

母音間の相対関係に基づく不特定話者母音系列の認識に関する研究

著者	下平 博
号	1112
発行年	1987
URL	http://hdl.handle.net/10097/9848

氏 名	しも 下 だいら 平 ひろし 博
授 与 学 位	工 学 博 士
学位授与年月日	昭和 63 年 3 月 25 日
学位授与の根拠法規	学位規則第 5 条第 1 項
研究科，専攻の名称	東北大学大学院工学研究科 (博士課程) 情報工学専攻
学 位 論 文 題 目	母音間の相対関係に基づく不特定話者母音系列の認識に関する研究
指 導 教 官	東北大学教授 木村 正行
論 文 審 査 委 員	東北大学教授 木村 正行 東北大学教授 城戸 健一 東北大学教授 竹田 宏 東北大学助教授 牧野 正三

論 文 内 容 要 旨

第 1 章 序 論

機械による連続音声認識の単位として音素を使用する場合，個人差，調音結合，セグメンテーション等の問題の解決が重要である。本研究は，これらの問題の中で個人差の問題に主眼を置き，日本語の認識において特に重要と考えられる母音を主な研究対象とした。

従来，母音の認識は予め用意してある母音テンプレートと未知入力母音を個別に照らし合わせて認識する手法（個別識別法）がとられてきた。個人差問題については，マルチ・テンプレート，正規化，適応化など多数の手法が提案されてきたが，未だ満足のおける結果が得られていない。

1984年に東北大学の倉によって発表された母音間距離における相対関係は，新しい音声認識手法の可能性を示したものである。個人差に強い相対関係を用いることによって，多数の母音を一つの系列としてとらえ，総合的に認識することが可能となった。本研究は，この相対関係をより深く分析することによって，相対関係法を一層発展させることを研究の目的とした。言い換えれば，音素の認識が従来，認識の対象とする音素区間のみに着目していたのに対して，単語あるいは文章のような広い区間に目を向け，全体の中の部分として認識する手法の開発を目指して本研究は始められた。

第 2 章 母音間距離による相対関係法の問題点

本章では，従来の母音認識法の問題点の提起を行うと共に，母音間の相対関係を利用して母音を

系列として認識する手法の原理を明らかにした。

まず、従来の代表的な母音認識手法を使って、本研究で使用する音声資料についての識別実験を行った。その結果、正規分布仮定によるベイズ識別や、マルチ・テンプレートの手法をもってしても、話者によって識別率に大きな差が生じることを示した。次に、個人性、調音結合の問題に対する立場を、従来の母音を個別に認識する手法（個別識別法）と相対関係法について述べ、両者の違いを明らかにした。

ついで、本研究で用いられる相対関係を利用した系列認識の手法（相対関係法）の原理について新たな説明を行なった。即ち、相対関係を利用した母音の系列の認識を最適候補系列として定式化を行ない、系列認識の手法を明確にした。本手法では、まず図1に示すように n 個の入力母音に対して予備識別を行い、それぞれ第3位までの候補を求めて候補リストを作成する。長さ n の入力母音系列に対する候補リストから得られる可能な母音候補系列は 3^n 通りある。任意の候補系列を

$$c_s = (c_{m_1}^{(1)}, c_{m_2}^{(2)}, \dots, c_{m_n}^{(n)}) \quad (m_p \in \{1, 2, 3\}, p = 1, \dots, n)$$

で表す。ここで、添え字 m_p は $c^{(p)}$ の第 m_p 要素（候補）を表す。いま、第 p 番目と q 番目の入力母音が母音組 $(c_{m_p}^{(p)}, c_{m_q}^{(q)})$ である確からしさ（2項関係の整合度）を $r_{m_p m_q}^{(p, q)}$ で表す。すると、候補系列 c_s の整合度 $R(c_s)$ は、全ての部分列の整合度の和として、

$$R(c_s) = \sum_{p=1}^{n-1} \sum_{q=p+1}^n r_{m_p m_q}^{(p, q)}$$

で与えられる。

$R(c_s)$ は純粹に相対関係に基づく整合度を与えるが、個別認識による母音個々の信頼度 $A(c_s)$ を使うことにより、候補系列の得点を $R(c_s)$ と $A(c_s)$ の重み付き線形和、

$$S(c_s) = \alpha R(c_s) + (1 - \alpha) A(c_s) \quad (0 \leq \alpha \leq 1)$$

によって定義することができる。従って求める最適候補系列 c_s^{opt} は $S(c_s)$ を最大にする系列として以下のように定義される。

$$c_s^{\text{opt}} = \arg \{ \max_{c_s} S(c_s) \}$$

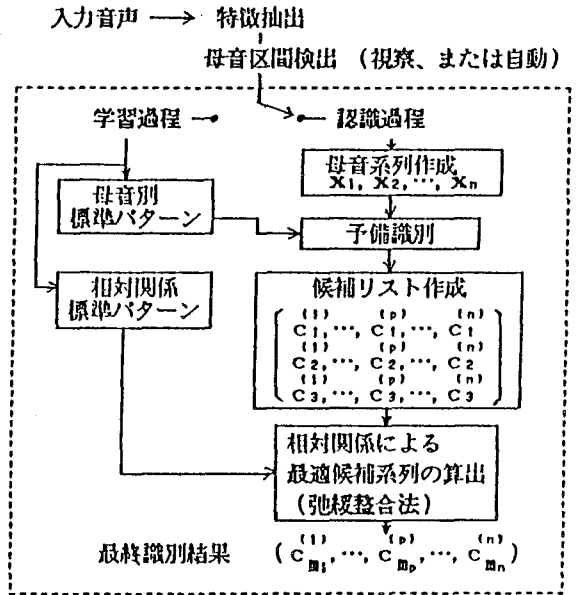


図1 相対関係法による母音系列の認識アルゴリズム

図1に示すように n 個の入力母音に対して予備識別を行い、それぞれ第3位までの候補を求めて候補リストを作成する。長さ n の入力母音系列に対する候補リストから得られる可能な母音候補系列は 3^n 通りある。任意の候補系列を

この最適候補系列を求めるには、枚挙法では非常に多くの計算時間を必要とする。高速に準最適解を求める手法として、確率的弛緩整合法を利用し、従来用いられてきたRosenfeldの修正公式より高速な手法の導入をはかった。

上記の準備の後、当初提案されていた母音間距離の相対関係に基づく母音系列の識別手法による識別実験を行った。その結果、系列に含まれる母音の種類に偏りがあると識別率が大幅に低下することが明らかになった。この原因について検討した結果、相対関係の定義方法に問題があることが判明した。これによって相対関係の評価には、話者に関する安定性の他に各相対関係間の識別性能による検討が重要であることを指摘した。

第3章 母音間の相対関係の定式化と分析

本章では新たに3種類の相対関係の定式化を行い、実験による有効性の分析を行なった。相対関係の有効性は、個人差に対する関係の安定性及び相対関係相互の識別性能の2点から分析を行なった。

定式化1は従来の母音間距離における相対関係を改良したもので、3母音間の関係（3項関係）を記述したものである。距離の相対関係の分布を事後確率に変換して分析した結果、分布が類似した相対関係が多いために、識別への寄与が非常に小さいことが分かった。このことは3母音組の識別実験によっても確かめられた。

定式化2は重回帰モデルによる母音推定を利用したものである。推定に用いる母音を1個とした（2項関係）モデルと、2個とした（3項関係）モデルを設定した。推定精度においては、3項関係モデルの方が2項関係モデルより優れていることが判明した。しかしながら、母音組の識別実験では、2項関係モデルの方が高い識別性能を示した。このことは3項関係モデルは組合せの数が多い分、識別が困難になっていることを意味し、推定精度と識別性能は必ずしも対応関係がとれるわけではないことを示している。

定式化3は2母音間の相対座標を新たな変量として、2項関係の整合度を正規分布仮定のベイズ識別に基づいて与えるものである。2母音組の識別実験では、定式化2よりも高い平均識別率を示すと共に、話者による変動も小さいことが分かった。

第4章 相対関係法による母音系列の認識実験

本章では、第3章で定義した3種類の相対関係の定式化に基づき、種々の母音系列による識別実験を行ない、それぞれの有効性を比較、検討した。

母音区間の検出を視察によって行った場合の実験では、3種類の安定化の中では、定式化3が識別率の点で最も優れており、次いで定式化2、定式化1の順であった。定式化3においては、安定な認識に必要とされる系列の長さは20程度あれば十分であることが分かった（図2）。単音節母音に関しては99.1%と非常に高い識別率が得られると共に、単語中の母音においても従来法より2%程度識別率が向上することが分かった。3章で示したように、個々の相対関係の識別精度は低いに関わらず、系列として総合評価すると高い識別率が得られる。このことは、系列中の数多くの

関係を使うことによって個々の曖昧性を減少させていることを意味している。

さらに、調音結合の影響を強く受ける連鎖母音に関する認識でも、従来法より高い識別性能を示し、本方式が個人差問題のみならず調音結合問題にも有効であることが判明した。また、誤識別の原因について解析し、弛緩法の誤りに起因するものは殆どないことを示した。

一方、母音区間を自動検出した場合の識別実験を行ない。検出誤りを含む母音系列に対しても、本手法が個別識別法よりも優位で、かつ安定した結果が得られることを示した。

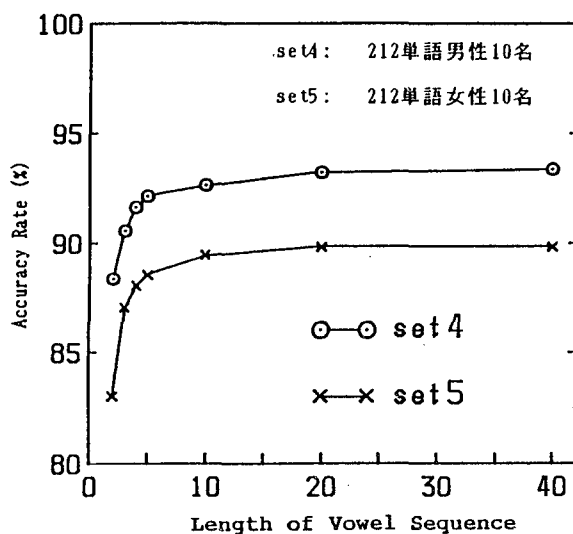


図2 系列長と識別率の関係（定式化3）

第5章 相対関係法の子音認識への適用

本章では、母音認識に使われてきた相対関係法を子音認識に適用する方法について検討した。時間・周波数スペクトルパターン(TSP)を基本として、従来のベイズ識別による個別認識法と本手法(相対関係法)との比較を行なった。後続母音独立の場合では、個別識別法に比べて相対関係法は識別率がかえって低下することが分かった。これは、後続母音の影響を強く受ける子音においては、相対関係にもその影響が強く現れ、かえって子音の識別に重要な情報が捉えられなくなっているためであると考えられる。一方、後続母音を1つの母音に限定した場合、相対関係法は有声破裂音で約3%、鼻子音で約5%、個別識別法よりも高い識別率が得られ、後続母音が同一であれば、相対関係法が有効であることが分かった。

第6章 結 論

結論として本論文を要約し、目標の達成度について述べた。

審 査 結 果 の 要 旨

単語あるいは文章中の母音を高精度で認識することは、音声の自動認識を実現するうえで重要な鍵となる。しかしながら、個人差や調音結合によって母音の音響的特性が複雑に変化するので、実用に耐え得るような母音認識の精度が得られていないのが現状である。著者は単語あるいは文章中の母音をそれぞれ個別に認識するのではなく、母音間の相対関係を利用する母音認識法を研究し、より高精度の母音認識を実現した。本論文はその成果をまとめたもので、全文6章よりなる。

第1章は序論である。第2章では、個別認識の手法及び相対関係を用いた従来の手法について、個人差及び調音結合の影響を考慮した種々の認識実験を行っている。その結果に基づいてこれらの手法の問題点を指摘するとともに、相対関係に基づく母音認識の原理を究明し、その厳密な定式化を与えている。これらは興味ある知見である。

第3章では、前章の結果を基にして、従来の母音間距離を用いた相対関係法を改善した定式化、母音推定を利用した定式化、さらに母音間の相対座標を利用した定式化を提案している。次に、それぞれの性能の分析と評価を行い、相対座標を用いた定式化が識別性能が高く、話者による影響も少ないと結論している。

第4章では、母音の系列長を変えた場合や種々の音韻環境での認識実験によって、相対関係法の性能評価及び従来の個別認識法との比較を詳細に行っている。その結果、相対関係法の認識率は母音の系列長が20程度で飽和する。単音節母音では認識率が99.1%に達する。単語中では個別認識法として定評のあるベイズ識別より約2%の精度の向上が見られる。検出誤りを含んだ母音系列に対しても有効であることなど、重要な知見が得られている。

第5章は相対関係法の子音認識への適用を試みたもので、子音の後続母音を限定した場合には子音認識率の向上が期待できることを示唆している。

第6章は結論である。

以上要するに本論文は、母音間距離を用いた従来の相対関係法の問題点を解明し、母音間の相対座標を用いた新たな相対関係を定式化することにより、種々の音韻環境においても安定に高い認識率が得られることを示すなど、音声認識の研究に有用な知見を加えたもので情報工学の発展に資するところが少なくない。

よって、本論文は工学博士の学位論文として合格と認める。